A Model for Synthetic Multivulva Gene Action

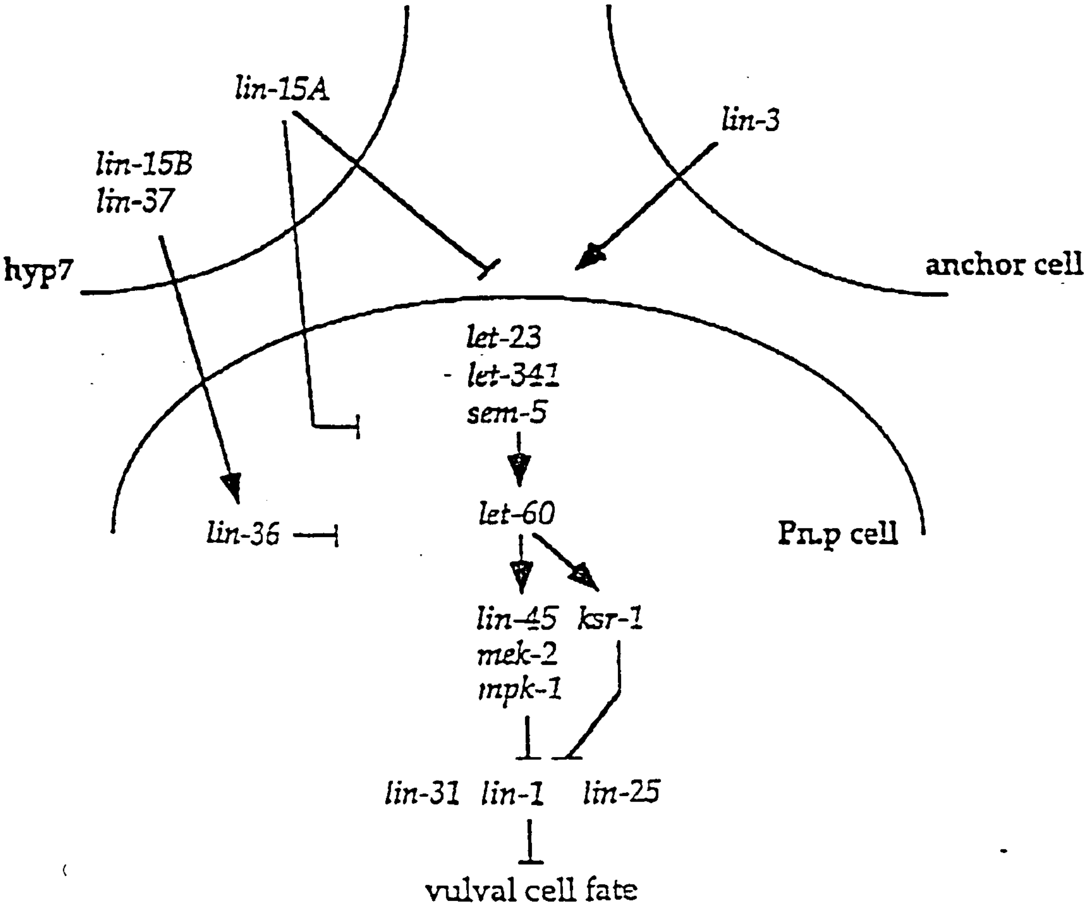

lin-15A
lin-3
lin-15B
lin-37
hyp7
anchor cell
let-23
let-341
sem-5
let-60
lin-36
Pn.p cell
lin-45
mek-2
mpk-1
ksr-1
lin-31
lin-1
lin-25
vulval cell fate

FIG. 1

>lin-37 protein sequence
MSEIDPLAEFLLPEDGDRNARQNDPLISGGPLPLESPSRKLTSLLSYDPTVPESPDMK
FARKRLGNLLTTIKHHPSEIIGVLPEDYTRADEEPGRQGRPPGRPRKMPRHESSTSLM
ESPRKTMTRDSKIMFELRGKPFEMIAGRFEEEYSLGRAWVKGHMNNEYEPIKAQRTD
YAPNLAVDYLACREIHRMPRPDKSIPELPIVPSRIDEFDATVDPRYETDLKNEYIRH
WKQVKKGWCAHQRRRTAPHARSIALINKIYQPGESKTVEQALGLI

SEQ ID NO:1

FIG. 2

>lin-37 cDNA sequence
ATGTCAGAAATAGATCCACTTGCCGAGTTCTTGCTTCCAGAAGACGGAGATCGAA
ATGCTCGTCAAAATGATCCATTGATAAGCGGAGGTCCACTTCCATTGGAATCGCCA
AGCAGAAAACTCACATCCCTGTTATCCTATGATCCGACAGTTCCGGAGTCACCGGA
TATGAAATTCGCCAGAAAACGTCTGGGAAATCTGCTGACAACCATAAAACATCA
CCCATCGGAAATAATTGGAGTACTCCCAGAAGATTATACTCGTGCTGATGAAGAG
CCCGGGCGCCAAGGACGTCCACCAGGTCGCCCTCGTAAGATGCCGCGTCACGAATCTT
CAACTTCACTTATGGAATCACCACGCAAGACTATGACTCGTGATTCTAAAATTAT
GTTTGAATTGCGTGGAAAACCATTCGAAATGATAGCTGGACGTTTTGAAGAAGA
ATATTCACTTGGTAGAGCATGGGTTAAAGGACACATGAATAATGAATATGAACC
AATAAAAGCTCAAAGGACAGACTATGCACCGAATCTGGCTGTTGATTATCTTGCA
TGTCGCGAGATTCATCGAATGCCACGTCCAGATAAATCAATTCCTGAGCTGCCAAT
TGTTCCATCTAGAATCGATGAATTCGACGCTACAGTCGATCCAAGATATGAAACA
GATTTGAAAAATGAATACATTCGTCATTGGAAACAAGTCAAAAAAGGTTGGTG
TGCTCATCAACGTCGTCGGACTGCTCCCCATGCAAGAAGCATAGCATTAATCAACA
AAATCTACCAGCCTGGAGAGTCGAAAACTGTCGAGCAAGCACTTGGTCTTATTTA
AATATTCTAACATGTAATTTCAATTTATCTCTTACTTTCTGATCTTGCTATCACA
TGTCTCTTATTTCAAAAATCTCACTTTAAAATTCATATAAATAATGGGTTTATT
CAAATACATCATCTTGAC

SEQ ID NO:2

FIG. 3

>LIN-35 protein sequence
MPKRAADEPGTSTTDPFHEQSPFDAVLAGTETTDTICEEPPAKRIDLDIKQEFNGGVQ
SGGLIKNESELTQMTIKQETEGNINEARREEEDEEQDEDSRTSMPPALGEDDDYEEDDA
AIYYYLLSKGIKRRGKTIRILIQPFPVSILTIANSFDISVAEMLDKTARFVEIIHSRKIR
RYQEYIRRIQEGLAVSCVIFKKFCRIFCKIFEEIKVGSENCPSSHELFTVLWTSFLVMK
SRMTVDDLISNYQLLFSILDQVYTEMCSMKEGIVHHLNQKFVEDLLENDCTIIRALCT
QFGGSVLDARHFSDHTFKKMEKTGIPSTWNFQEFRDLIMNVPKTAYENYLLQRGSID
ERIFIPSVEDFSKIFQSPDTYSVADILKVSYSGRRFRDAEFLTKISNNHCLEKLALGGK
VASEKLVTQSKEQPRVPCVEYNLELGNYPDDLESNNQSLYNRLTKIIGSWKLENSKLE
EVCGTMSDSPMATILLKSDEMTNKFERTLSAELGETINENIPKYHYNVRKELELVFLI
FMEKIIVAELKKKVREEDLLNVIRREEFLDSVFCFCVELILVSNGYDRPFPWSAELCG
VHPFMFHKVIDLMITHEKQLSRQMVQHFSRIEETVIEYFSWKSDSPLWPMVVRCPF
AHFQEFGEDWADKLNSYSPIKFTPIKKPDDLRDELGRPIVPQNQTSRTLRIFLKRTYFT
AARRLQDLTDRVSMGARAKSQCWSLFDYLLRNDTLIFMDRHLDQILLCCVFVIMKI
NESSMLFTEIMAQYRRQSANSLLVYRSVTVFQEQLNPENPQAVNTKETILERLEGPQ
KEKTTVDIIKYYNIEFRDRIKYIIGQIDSASDEDLMEMPVATESGLMPVRVYLTHKLS
IQTLPKTKHGESKQERAIANLEKSGITIAMERSGD **SEQ ID NO:3**

FIG. 4

>lin-35 cDNA sequence
CCAAGTTTGAGACTGGTAAAAACTTCTTCAATATGCCGAAACGAGCAGCCGATGA
GCCTGGAACATCAACAACTGACCCATTTCACGAGCAAAGCCCATTCGATGCCGTGT
TAGCCGGCACGGAGACAACGGATACAATATGTGAAGAGCCACCAGCAAAACGAA
TCGACTTAGATATAAAGCAAGAATTCAATGGTGGAGTGCAAAGTGGAGGGCTGA
TTAAAAATGAATCCGAATTGACTCAAATGACAATCAAACAAGAAACAGAAGG
AAACATAAATGAAGCTAGACGAGAAGAAGAAGACGAAGAACAAGATGAAGA
TTCCAGAACATCAATGCCACCTGCATTGGGAGAAGATGATGATTATGAGGAGGAT
GATGCTGATAGTTTTATTGATAAAACTAATACACCGCCACCATCACAATCATTTC
TGGAAGGATGTCGAGCAGCTAATTTACCAAATGACATTGTTACTGGTGCATGGGA
AACGTACAACCACGCGGTTCAACGGGTTTCTCTTGAGGGTTCGGAATCGGCGTGGCA
ACTATCAGCAATTTACTATTATCTTCTATCAAAAGGAATAAAACGTCGTGGAAA
AACAATCCGTATTCTCATTCAACCGTTTCCTGTTTCGATACTTACAATTGCCAACT
CATTTGACATATCCGTTGCTGAAATGCTTGACAAAACTGCTCGATTTGTGGAAAT
TATACATTCCAGAAAAATTCGTCGTTATCAAGAATATATTCGACGAATTCAAGA
AGGACTCGCAGTTTCTTGTGTGATATTCAAAAAGTTTTGCCGAATTTTCTGCAAA
ATATTCGAGGAGATCAAAGTTGGATCCGAAAATTGTCCATCTTCTCATGAACTTT
TTACGGTTCTTTGGACATCTTTTCTGGTGATGAAAAGTCGAATGACAGTGGACGAT
TTGATTTCAAATTATCAACTTCTT-TTTTCAATACTTGATCAAGTATATACCGAA
ATGTGTTCAATGAAAGAGGGAATAGTTCATCATTTGAATCAAAAATTTGTTGA
AGATCTTCTTGAAAATGATTGTACGATTATTCGAGCTCTTTGCACACAATTTGGT
GGAAGTGTTCTTGATGCACGGCACTTTTCTGATCATACTTTTAAGAAAATGGAGA
AGACTGGAATTCCGTCCACTTGGAATTTTCAAGAGTTTCGAGATTTGATCATGAA
CGTTCCAAAAACGGCATATGAGAATTATCTATTGCAACGTGGAAGTATTGATGA
GCGGATTTTCATTCCAAGCGTTGAGGACTTTTCAAAAATTTTCCAATCCCCGGACA
CATACTCAGTAGCAGATATTCTCAAAGTGTCTTACTCTGGAAGACGTTTCCGTGAT
GCAGAATTTCTTACAAAAATCTCAAATAATCATTGTCTGGAAAAGTTGGCATTA
GGTGGAAAAGTAGCATCAGAAAAGTTGGTAACACAGTCAAAAGAACAGCCGAG
AGTTCCGTGTGTTGAGTATAATCTCGAATTGGGAAATTATCCAGACGATTTGGAA
TCGAACAATCAAAGTCTTTATAATAGATTGACAAAAATTATTGGAAGCTGGAA
ATTGGAGAATTCGAAACTCGAAGAAGTGTGTGGCACAATGTCCGACAGTCCAATG
GCAACAATTCTTCTGAAAAGTGATGAAATGACAAATAAATTCGAGCGAACTTT
ATCTGCAGAACTCGGAGAGACGATCAATGAGAATATTCCTAAATATCACTATAA
TGTTCGAAAAGAATTGGAATTAGTTTTTCTCATTTTCATGGAGAAAATAATTGT
TGCAGAATTGAAAAAGAAAGTACGAGAGGAGGACTTGCTGAATGTGATTCGTCG
GGAAGAATTTCTTGATTCAGTTTTCTGTTTCTGTGTTGAACTGATCCTTGTTTCCA
ATGGATATGATCGTCCATTTCCATGGAGTGCTGAACTGTGTGGAGTACATCCATTT
ATGTTTCATAAAGTAATTGATTTGATGATAACACATGAGAAACAGCTAAGTCGT
CAAATGGTTCAACATTTTAGTCGAATTGAAGAAACTGTAATTGAGTATTTTTCG
TGGAAGTCTGATAGTCCATTATGGCCAATGGTTGTCAGGTGTCCATTTGCACATTT
TCAAGAATTCGGAGAGGATTGGGCTGATAAATTAAACTCGTACTCACCAATAAA
ATTCACTCCAATCAAGAAGCCTGATGATCTACGAGACGAACTTGGAAGACCTATA
GTTCCTCAAAATCAAACTTCAAGAACTCTAAGAATTTTTTGAAAAGAACTTAT

FIG. 5 **PAGE 1 OF 2**

TTCACCGCCGCTCGTCGACTTCAAGATCTCACTGATCGTGTTTCAATGGGAGCTCGTG
CAAAATCACAATGCTGGTCACTTTTCGATTATCTTCTTCGCAATGACACTTTGATT
TTTATGGATAGACATCTTGATCAAATTCTTCTTTGTTGCGTGTTTGTCATTATGAA
GATAAATGAGTCATCAATGCTTTTCACGGAAATAATGGCTCAATATCGACGACA
ATCAGCCAATTCTTTGCTGGTCTACCGAAGTGTTACAGTATTCCAAGAACAACTG
AATCCCGAAAATCCACAGGCAGTAAACACGAAGGAGACAATTTTGGAACGTCTT
GAAGGTCCACAAAAAGAAAAAACGACAGTTGATATAATCAAATATTATAATA
TCGAGTTTCGGGATCGTATCAAGTATATTATCGGTCAAATTGATAGTGCTTCAGA
TGAAGATTTGATGGAAATGCCGGTTGCAACAGAATCTGGATTGATGCCTGTTCGA
GTTTATTTAACACATAAATTATCGATTCAAACGCTTCCAAAAACGAAACACGGA
GAGTCGAAACAAGAAAGAGCTATTGCGAACCTTGAAAAATCTGGAATTACGATC
GCTATGGAACGGTCTGGAGATTAAAAATGATTGTTGTGAATACTTTGAACTTTTT
AATGCATTTTTGATTAATCATTTAGTACTTCTTTTCTCGTCTATTTTTTTATCTTT
TCCTTCAAATTCAGGCAAGTAATTATACTTTCCATTTCTAATTGATTGCTTCAAA
ATAGACGTCTAGTTATATTCAAAACAATCCCCCTTTTGAATTGGAATCTTCAAAT
ATCGTATTAAATATTAATATTGTAATCATTTTTCACAATCCCCCATGCCATTATT
GTTACTGATTTTTTCTCTCTTTTTAACCATCATCGATAAATTCATTTTACAGTTAT
AAAAAAAAAAAAAAA

SEQ ID NO:4

FIG. 5 **PAGE 2 OF 2**

>LIN-53 protein
MATLEDGTSEDRVANDEYKIWKKNTPFLYDLVMTHALEWPSLSVQWLPDVAKDN
SDHTIHRLILGTHTSDEQNHLLISKICMPTDDAQFDASRYDTERSEYGGFGAVNGKVE
PDIRINHEGEVNRARYMPQKSNIIATKSPHADVYIFDYLKHSAVPRDNTFNPLIRLK
GHTKEGYGLSWNPNKEGLILSASDDQTVCHWDINANQNVAGELQAKDVFKGHESV
VEDVAWHVLHDGVFGSVGDDKKLLIWDVRTSTPGHCIDAHSAEVNCLAFNPYSEFI
LATGSADKTVALWDLRNLRMKLHSFESHRDEIFQVQWSPHNETILASSGTDKRLHV
WDLSKIGEDQSAEDAEDGPPELLFIHGGHTAKISDFSWNPNEPWVVCSVSEDNILQV
WQMADNIYNEVDEETPADVVERQQ **SEQ ID NO:5**

FIG. 6

>lin-53 cDNA
GAAATGGCCACTCTTGAAGATGGAACCTCCGAAGATCGCGTCGCAAATGATGAAT
ACAAAATTTGGAAAAAGAACACCCCGTTCTTGTACGATCTCGTAATGACACATGC
GCTTGAGTGGCCTTCACTCAGTGTTCAATGGCTCCCAGACGTTGCGAAGGATAACAG
CGACCATACTATTCATCGGCTCATTCTTGGAACTCATACTTCAGATGAGCAGAATC
ACTTGCTTATTTCTAAGATCTGTATGCCAACGGATGATGCCCAATTTGATGCATCT
CGCTACGATACCGAGCGCAGTGAATACGGTGGTTTCGGAGCTGTTAACGGAAAAGT
GGAACCCGATATTCGCATTAACCACGAAGGGGAGGTTAACAGAGCTCGTTACATG
CCTCAAAAGTCGAATATCATTGCTACAAAGTCTCCACATGCTGATGTTTACATTT
TCGACTATTTAAAGCACTCTGCTGTTCCTCGTGATAACACGTTCAATCCGCTTATCA
GACTGAAAGGACACACGAAGGAAGGCTATGGATTATCATGGAATCCAAACAAA
GAAGGTTTGATTCTATCAGCGTCAGATGATCAGACAGTTTGTCATTGGGATATCA
ACGCAAATCAGAATGTTGCCGGGGAATTGCAAGCGAAGGATGTTTTCAAAGGTCA
CGAGTCAGTCGTTGAAGATGTTGCTTGGCACGTTTTGCATGATGGTGTCTTCGGATC
GGTTGGTGACGATAAGAAATTGCTCATTTGGGATGTGCGCACAAGCACTCCTGGAC
ACTGTATCGATGCTCATTCTGCCGAAGTTAACTGTCTCGCATTCAATCCATATTCCG
AATTCATTCTGGCCACCGGATCAGCTGATAAAACTGTCGCTCTTTGGGATCTACGT
AATCTACGAATGAAACTTCACTCATTTGAATCACATCGTGATGAAATCTTCCAAG
TTCAGTGGAGTCCACACAACGAGACTATTCTTGCATCCAGCGGTACTGATAAACGT
CTTCATGTGTGGGACCTATCTAAGATTGGAGAAGACCAATCTGCCGAAGACGCGGA
AGATGGTCCACCAGAGCTGTTGTTTATTCACGGTGGGCACACCGCCAAGATCAGCG
ATTTCTCGTGGAACCCGAACGAGCCTTGGGTTGTGTGCAGTGTGTCAGAAGACAAT
ATTCTCCAAGTGTGGCAAATGGCTGATAACATATACAACGAAGTTGACGAAGAA
ACTCCAGCCGATGTGGTAGAGAGACAACAGTAAAATACGTGAAACGCGCTTAAA
TTATTTGTATTTAACTTCTATCCTTCTTTAATTTTGCATCTCAACAAATTGTTCAT
CTTACCATTTATTCAAACGCATATTCTTCACCAACTAAGTTTTTAAAGTTAAAA
TGTTACCTTGAGATATGATCATATTTTGTTGAACCTGAAATAAATTCGATGACCA
TTGTCAAAAAAAAAAAAAAAAA **SEQ ID NO:6**

FIG. 7

PARTIAL *lin-55* TRANSCRIPT WITH PREDICTED TRANSLATION

```
5' aaaatcttga ttcgtccagg caggtcctgt agaaatgaac ccgacaaact acgatccccg tatcggccaa ccggcacaaa gtcgaccaca agtatacccct 100

gagtATGGGA AGACGAATTG TTCAAATGCC TACAGGTCTT CCTCGTTCAT ACCAAGATGA GAGTCACAAT GAACCTGTGG GCTGGGATGA ACCTTCGGGA 200
    M G   R R I V   Q M P   T G L   P R S Y   Q D E   S H N   E P V G   W D E   P S G
GTTGGAGGAT CTAGTGGAGC AGGAGGTCAA CAATCTGACA AACCAACCGG ACTCAGACAT TTTTCCACGA AAGTTTGTGA AAAGGTGAAA GAAAAAGGAT 300
V G G S   S G A   G G Q   Q S D K   P T G   L R H   F S T K   V C E   K V K   E K G L
TGACGAATTA CAATGAAGTG GCTGATGAAC TGGTAGCAGA TTACTTTCAA AATAATCTTA TCAAGCAGAT TGATGTTGTG AAACAAGAAT ACGATATGAA 400
  T N Y   N E V   A D E L   V A D   Y F Q   N N L I   K Q I   D V V   K Q E Y   D M K
AAATATTCGG AGGAGAGTAT ACGATGCGTT GAACGTTCTT TTGGCAATGA ATATCATCAC AAAAAGCAAG AAAGATATCA GATGGATTGG ACTTCCTGCT 500
 N I R   R R V Y   D A L   N V L   L A M N   I I T   K S K   K D I R   W I G   L P A
TCAGCTTCAC AAGAAATTTC GAGATTGGAA GAAGAGAAAT CGCGACGGGA AGCCAGCATA AGTTCCAAAA AGCAGGCTCT CGAAGAAATG GTATTGCAAA 600
S A S Q   E I S   R L E   E E K S   R R E   A S I   S S K K   Q A L   E E M   V L Q I
TTGTTTCTTA CAAAAATTTG GTCGAACGGA ATAGAAAAAA TGAGCACAAA AATGGTCGCC CAGAAAACGA CACTGTACTT CATCTTCCGT TTTTAATAAT 700
  V S Y   K N L   V E R N   R K N   E H K   N G R P   E N D   T V L   H L P F   L I I
TAATACTGAT AAAGAAGCAA ATGTGGAATG CAGTGTATCA TCAGACAAAT CCGAATTTCT GTTCTCGTTC GACAAGAAGT TTGAAATTCA CGATGATTTC 800
 N T D   K E A N   V E C   S V S   S D K S   E F L   F S F   D K K F   E I H   D D F
GAAATTTTAA AAAAGTTGAA CCTTGCATGC AGTTTGGAGA CTACAAATCC CACAGCGGAA GAAGTCAAAA CTGCGAAAAG CTTTTTGCCA ACCTTACACC 900   SEQ ID NO:7
E I L K   K L N   L A C   S L E T   T N P   T A E   E V K T   A K S   F L P   T L H Q
AGCACTATGT CGACGAAATC ATTGCCAACC GTAAAAAGGT TGAAGCAGAA AAAGAGGAAA AGAGAAAACA ACAACAGTTG ATTGCTGATC AAATGTCTAT1000
  H Y V   D E I   I A N R   K K V   E A E   K E E K   R K Q   Q Q L   I A D Q   M S M
GAACCTCTCA CAAGCCCAAT ACGTTGAGCC GACAAGCTCT TTGGCGCAAA TGTCATATTC ATCGCGATTC AACAGACAGC TTCAGGAGCA TTTGATAAAC1100
 N L S   Q A Q Y   V E P   T S S   L A Q M   S Y S   S R F   N R Q L   Q E H   L I N
GATGGAAGTG AGGACAGAAG TGCTGCTGCC GGTATAATGG AACGCGATTA TGATATGGAT AAGAATGTTA ATCAGGGATC AGCCAGCCGA GGGCCAATGT1200
D G S E   D R S   A A A   G I M E   R D Y   D M D   K N V N   Q G S   A S R   G P M Y
ATAATACTTA TAGTCCACAA AAGATTCGAG CTCAAGTTAA CACGCCTCTT CAAGTACCAC CCGTTACCAA ACGATATTAC GTACAGAAGA CACAAGGACC1300
  N T Y   S P Q   K I R A   Q V N   T P L   Q V P P   V T K   R Y Y   V Q K T   Q G P
AATGAAGCAT GATATGACTC CCGTCGTTCG AACTGTCAAT CGGCCGTACT CTACTGTACC TCCAGATCGC AGGCTCTCAA CTGGAGCTAC TTCTGTGAAT1400
 M K H   D M T P   V V R   T V N   R P Y S   T V P   P D R   R L S T   G A T   S V N
TCAGGACCTG TAAAATACTA CGTGCCACAA GGACATCAAC CGATGCATCA GCCAGTTGGT CAGAGATATA GAGTTCGTCC ACAGCAACCA CAAATGAGTC1500
S G P V   K Y Y   V P Q   G H Q P   M H Q   P V G   Q R Y R   V R P   Q Q P   Q M S H
ATATGGGTCA GCCACATCAA GTGCAACAGA GAGTTGTCTA TCCTGCTGGA AGCATTTCTG GACATCAGTT ACAACCTGGA CAACGAATCG TAACTCAGCC1600
 M G Q   P H Q   V Q Q R   V V Y   P A G   S I S G   H Q L   Q P G   Q R I V   T Q R
AATTGTTGCT CCAGGTGGTC CACACCCGCC GGGCACAATT GTTCGGAAAG TGATTCGTAA AATTGTTGTC AACAATCCAA AGCAATCTCC AGCTCAACAA1700
I V A   P G G P   H P P   G T I   V R K V   I R K   I V V   N N P K   Q S P   A Q Q
GTTATACAAA AGAAAATGAT GGAGCAAGAT ATGTGCACAT TTGAACGCAA AACGGAACAG CCGATGACTT CTGCGCAGGC CGCTGCTCTC ATTCAACACC1800
V I Q K   K M M   E Q D   M C T F   E R K   T E Q   P M T S   A Q A   A A L   I Q H P
CTCAACCGGA GGAATACGAT TATTTCCAGT AA  3'                                                                  1832   SEQ ID NO:8
 Q P E   E Y D   Y F Q .
```

Notes: The 3' untranslated region has not been sequenced.

FIG. 8

C.elegans **E2F-1 Protein Sequence**

1
MELQKALEMTKQSSIKNNLMLGLDNELDFDFDFDEDEDLDQPQMGTR
ADKSLGLLAKRFIRMIQYSPYGRCDLNTAAEALNVRQKRRIYDITNVLE
GIGLIEKRSKNMIQWKGGDFMLNVKEGKRLSATTEEEDRMEQLKAEIE
QLNKEEELIEQRQRWLQQSLRNMTESVENNKLSYVLRSQLAEIQGSDLT
IGIQTRVGTQVRLSDPEQVEIHGGPSWCYLKDPSGPLRAAIVSNHELHDF
VQRERAKRPGEEHVDADAPDEMMDDSRYRNRRTINDDEMFGFEQKVP
AMKHLEPPPASDDYVYSSTGDEYRGDSIVDLYGD
324 **SEQ ID NO:9**

FIG. 9

C.elegans E2F-1 cDNA Sequence

1

GTTTGAGCCATGGAAGACTCGTACAACGACATGGAAGACCCCGGC
TTCCGCCAATTATCTGACATGGAGCTTCAAAAAGCGCTGGAAATGA
CCAAACAGAGCTCGATAAAGAACAATTTGATGCTCGGGCTCGACA
ATGAGCTTGACTTTGATTTTGATTTTGACGAGGATGAGGACCTGGAT
CAACCACAAATGGGCACACGAGCCGATAAATCGTTGGGATTGTTG
GCGAAACGATTTATTCGAATGATTCAGTACTCACCGTATGGAAGAT
GCGATTTGAACACTGCCGCCGAGGCGCTCAATGTCCGGCAAAAGC
GACGAATCTACGATATTACGAATGTTCTCGAAGGAATTGGTCTTATT
GAGAAAAGAAGCAAGAATATGATACAGTGGAAAGGCGGTGATTTT
ATGCTAAACGTGAAGGAAGGGAAACGACTATCGGCCACAACAGA
AGAAGAAGATCGAATGGAACAATTAAAAGCTGAAATTGAGCAATT
AAATAAGGAAGAAGAGCTCATTGAGCAACGTCAAAGATGGCTTCA
GCAGAGCCTCCGAAACATGACAGAATCCGTGGAGAACAACAAGCT
CAGCTATGTGCTCCGTTCACAGCTCGCCGAGATTCAAGGCTCAGAT
CTTACGATTGGAATTCAAACGAGAGTCGGCACACAAGTTCGGCTCA
GTGATCCGGAGCAAGTCGAGATACACGGTGGACCATCTTGGTGTTA
CCTGAAAGATCCCTCTGGACCCCTCCGAGCCGCCATCGTTTCCAAC
CATGAGCTACATGATTTTGTACAGAGAGAACGAGCAAAACGGCCT
GGTGAAGAGCACGTTGACGCTGATGCTCCAGATGAAATGATGGAT
GATTCAAGATATCGGAATCGGCGGACGATCAATGATGATGAAATG
TTTGGTTTTGAGCAGAAAGTCCCAGCGATGAAGCATCTGGAGCCAC
CACCGGCCAGCGATGACTACGTTTATTCGAGCACCGGAGACGAGT
ATCGAGGAGATTCTATAGTCGATTTGTACGGAGATTAATTATTTTAA
TATTTTTTTTTTAAATTTCGAATTCTGCGACCATTTCTCATTTGACATC
TATTCATTTACTCCAAATTCCAAATTTTTCCCCAAAAAAATTATCGA
TGTTTCGGCTCCAAATGTTATTATTTTCCCATCCACAGTGCCCACAC
AATTCATAATGTGCCTCTGGAGAAAACCTAACGTATTTCAATTTCTA
TCCCAAATTTTTATTTTTCAAAAATTTCTCAGATTTTTAAATTATTTG
TCACACTTTTTTCTGTATTCAAACTGAACTTTTTCACTTGGATTTGTA
CGTTTTTTTTTTTGTTCAATTTTAATGGATTTTCACTTGAAAACCCCA
ATAAAAACGGGATAAATCGACGTTTTTGAATAAAAAAAAAAAAAAA
AA

1428

SEQ ID NO:10

FIG. 10

LIN-52 Protein Sequence

1
MSRPLGFIGYEFGDDEMFVQQMIEKKSNAEQAKMLEQQKKMLECTET
MPEESEPVPMKCLDFEEAFQSESVSKGYESPYKNISFLKEDAVTVNTMSH
CPADDIAKLIRNIQNSVYTLGIEEARQCRRGKLLNVLKPTGSASPRYLQP
TPPKNVAEETTGSQ
161 **SEQ ID NO:11**

FIG. 11

lin-52 cDNA Sequence

1
ATGTCGCGTCCGCTAGGATTTATTGGATACGAATTTGGAGATGACG
AAATGTTCGTCCAACAAATGATCGAAAAGAAATCAAACGCAGAAC
AGGCGAAAATGCTTGAACAACAGAAAAAGATGCTCGAATGCACCG
AAACAATGCCAGAAGAAAGTGAGCCAGTTCCAATGAAATGTCTCG
ATTTTGAAGAAGCATTTCAAAGCGAATCAGTATCAAAAGGTTACGA
ATCGCCATACAAGAATATTTCGTTTCTCAAGGAAGATGCTGTGACT
GTTAATACAATGAGCCACTGCCCAGCCGACGATATCGCCAAGCTCA
TCCGAAACATTCAAAACTCGGTGTACACTCTTGGAATCGAAGAAGC
TCGCCAGTGCCGACGTGGAAAGTTGCTCAACGTGCTGAAACCCACT
GGCTCGGCTTCTCCGAGATATTTGCAGCCAACACCACCGAAAAATG
TAGCGGAAGAAACGACAGGAAGCCAGTGAAATTGAA **SEQ ID NO:12**
493

FIG. 12

lin-54 Protein Sequence

1
MNQGEIVYQDDDDYYDESEIYDNYEEGAEFIEVNGQLVPHNPNLQAQ
QNRPGTSSMIQQHNRSMEVNQGLVKDEPIDTSSHRVYVPPPRPVQRKL
WKLFQPGPSTPGSSQYTVRNLSNLSGSPSMYDRQPASLPRTVQPMGLEM
GNSEQRKVYIDMKDHVSHIRLKTKKKVFAPGQRKPCNCTKSQCLKLYC
DCFANGEFCRDCNCKDCHNNIEYDSQRSKAIRQSLERNPNAFKPKIGIA
RGGITDIERLHQKGCHCKKSGCLKNYCECYEAKVPCTDRCKCKGCQNT
ETYRMTRYKNSGGAVSNTNALMSLTNASSTATPDSGPGSVVTDEHGDD
YEDMLLSHKPKVEMDPRRFPWYYMTDEVVEAATMCMVAQAEEALNY
EKVQTEDEKLINMEKLVLREFGRCLEQMITNTTELTQDLDAAPTDDIPG
PSTSTS
438

SEQ ID NO:13

FIG. 13

lin-54 cDNA Sequence

1
ATTTTCAGTGTTGACAATGAATCAAGGAGAAATCGTTTATCAAGAC
GACGATGATTATTACGACGAATCGGAGATATACGATAATTATGAAG
AAGGTGCCGAATTTATCGAAGTTAATGGACAGCTTGTGCCTCATAA
TCCAAACTTACAGGCGCAGCAAAATCGTCCGGGAACCTCGAGTAT
GATTCAACAGCATAATCGATCAATGGAAGTTAATCAGGGATTGGTC
AAAGACGAACCAATTGATACATCATCGCATCGCGTCTACGTCCCCC
CTCCGAGACCAGTTCAGCGAAAACTTTGGAAGCTTTTTCAGCCTGG
GCCCAGCACTCCCGGATCGTCTCAGTACACTGTGCGGAATTTGTCC
AATTTATCGGGTTCACCTTCAATGTACGATCGACAGCCCGCTTCATT
ACCTAGAACAGTGCAACCAATGGGCTTGGAGATGGGAAATTCTGA
ACAGCGAAAAGTTTACATCGATATGAAAGATCACGTTAGTCATATT
AGATTGAAAACTAAAAAAAAAGTATTTGCACCTGGCCAGCGGAAA
CCATGCAATTGCACGAAATCTCAATGCCTCAAGCTCTACTGTGATT
GTTTCGCCAATGGAGAGTTCTGTCGTGACTGCAATTGCAAGGATTGT
CACAATAATATAGAATACGACAGTCAGCGTTCAAAAGCCATCCGT
CAGTCACTTGAGCGAAATCCGAACGCTTTCAAGCCAAAAATTGGTA
TTGCTCGTGGAGGTATTACCGACATCGAACGTCTTCATCAGAAAGG
ATGTCACTGTAAAAAGAGTGGTTGTCTGAAAAACTATTGTGAGTGT
TATGAAGCAAAGGTTCCGTGTACCGATCGATGCAAGTGCAAAGGA
TGTCAGAATACTGAAACATACAGAATGACAAGATACAAGAACTCC
GGTGGTGCCGTGTCCAATACGAATGCCCTGATGTCATTGACCAACG
CTTCCAGCACAGCGACTCCAGATTCTGGTCCGGGAAGTGTGGTGAC
CGATGAGCATGGAGACGACTACGAGGATATGCTTCTTTCGCATAAA
CCGAAGGTCGAGATGGATCCTAGACGCTTCCCGTGGTACTATATGA
CCGATGAAGTCGTTGAGGCAGCCACTATGTGCATGGTTGCTCAAGC
TGAAGAAGCTTTAAACTACGAAAAAGTGCAAACCGAAGACGAAA
AACTCATCAATATGGAGAAGCTTGTCCTTCGTGAATTCGGCCGCTGT
CTCGAACAAATGATCACAAACACAACTGAGCTCACACAAGATCTT
GATGCCGCTCCAACGGATGACATCCCAGGACCATCTACTAGTACTT
CTTAACTATTCGCATTAAAATTATTATCAATTTTATCACAGTTGCGC
GATCTTTTATGATCTCACCTCTCACACATTCTTTGCCTTCCTCCCCTC
CTCTCAATGCTTTTACAGATTCACAAGTTGCCTTCTTTCAAAGTTGTC
AAATAAAAAAATGATCAGAAAAATTTGTTTCAT
1503

SEQ ID NO:14

FIG. 14

lin-52

lin-52 maps on LGIII between *unc-16* and *unc-69*.

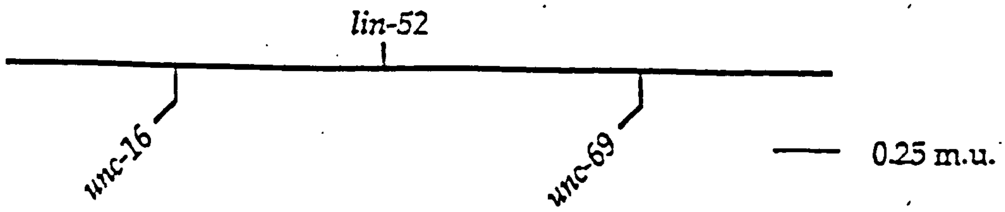

lin-52
unc-16
unc-69
0.25 m.u.

FIG. 15

lin-55

lin-55 maps on LGII between *rol-6* and *unc-4*.

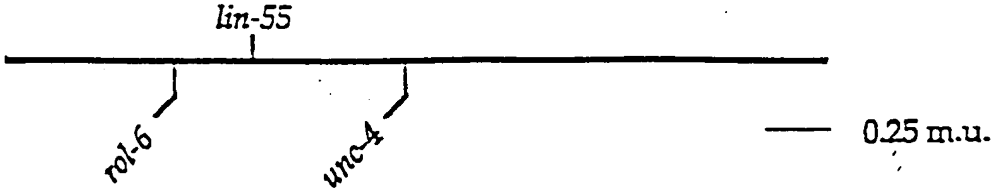

lin-55
rol-6
unc-4
0.25 m.u.

FIG. 16

lin-37 **message is present in both embryonic and mixed-staged RNAs**

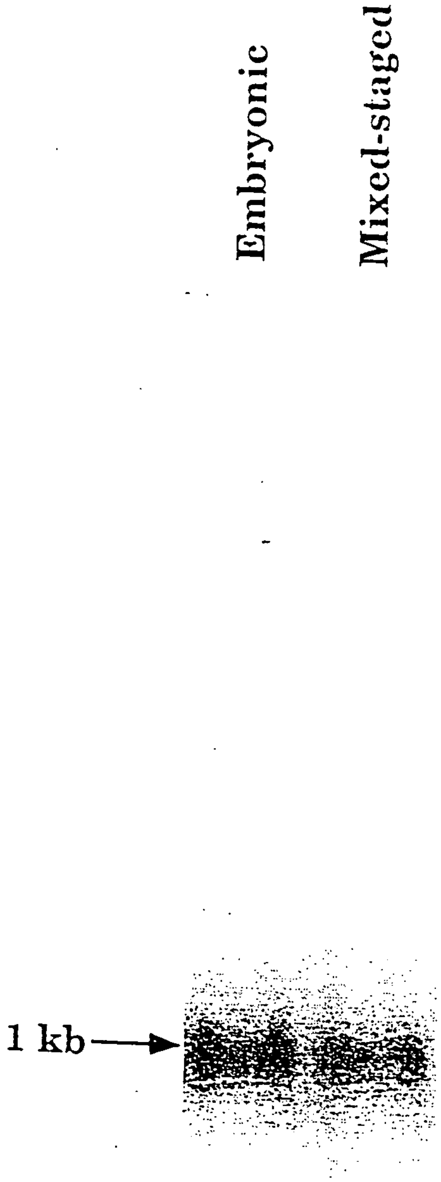
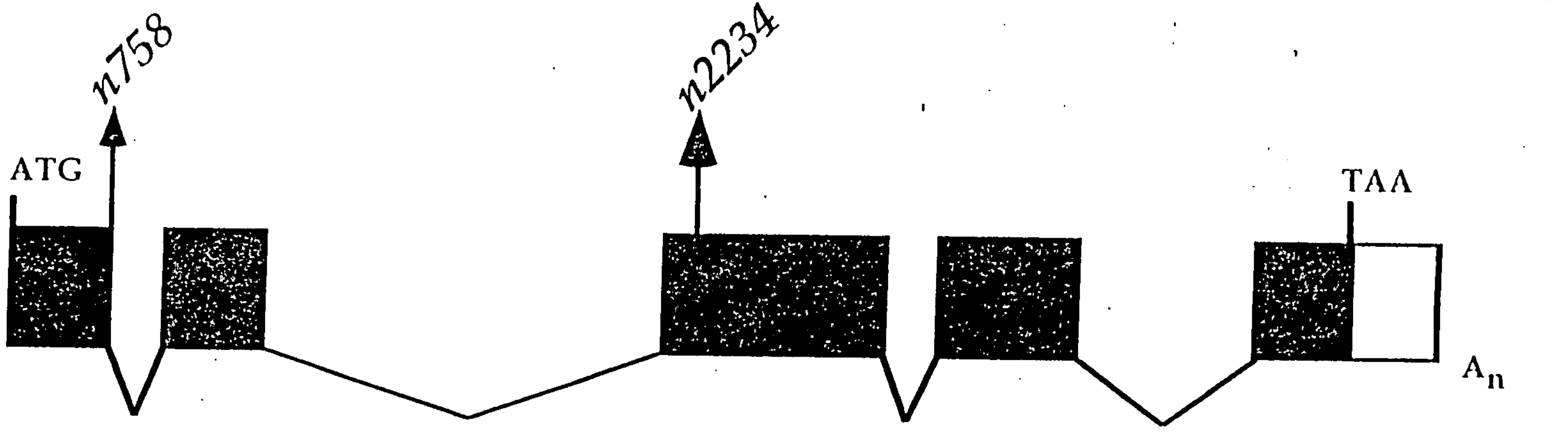

Embryonic
Mixed-staged
1 kb

FIG. 17

lin-37 Gene Structure and Mutations


500bp
ATG
n758
n2234
TAA
A_n

n758 G → A, splice donor

n2234 AAG → TAG, K105amber

FIG. 18

lin-37::GFP

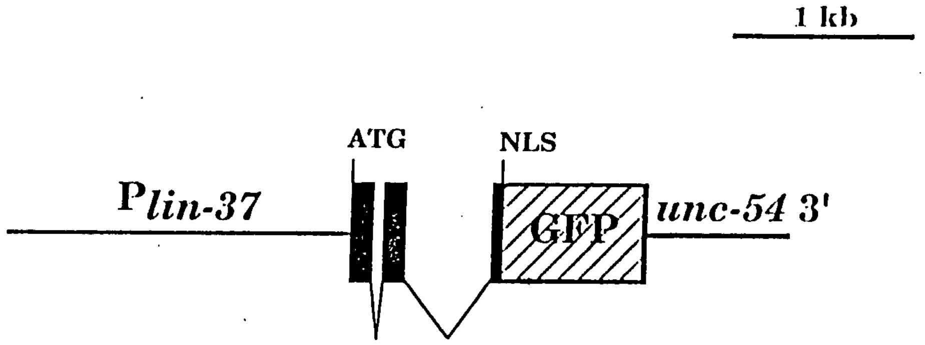

1 kb
ATG
NLS
P*lin-37*
GFP
unc-54 3'

FIG. 19

lin-37 Rescue

The *lin-37* minimal rescuing fragment contains five ORFS as predicted by Genefinder. Only the ORF C.cand2 is required for rescuing activity, therefore is likely to be *lin-37*.

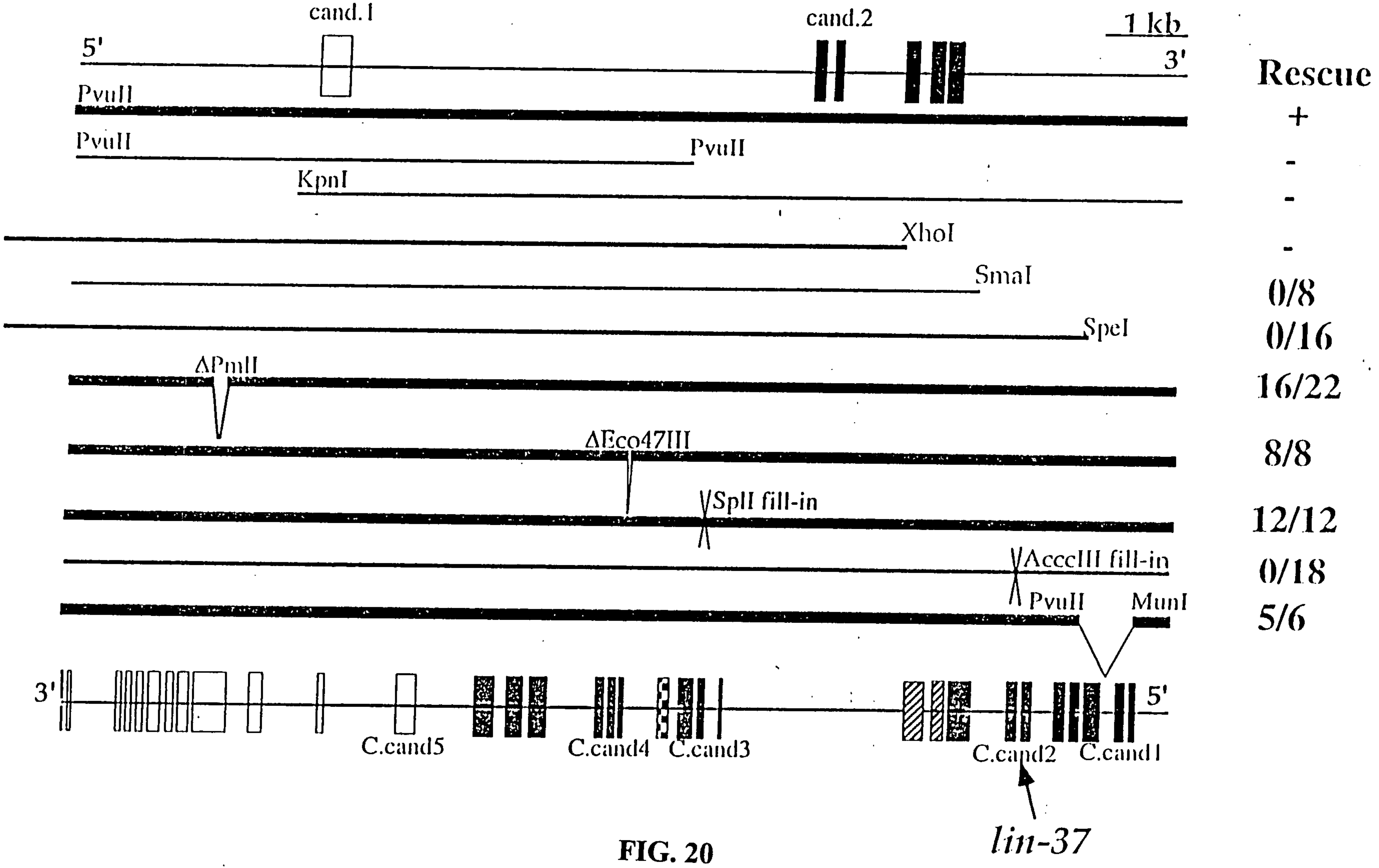

cand.1
cand.2
1 kb
5'
3'
Rescue
PvuII
+
PvuII
PvuII
-
KpnI
-
XhoI
-
SmaI
0/8
SpeI
0/16
ΔPmlI
16/22
ΔEco47III
8/8
SplI fill-in
12/12
AccIII fill-in
0/18
PvuII
MunI
5/6
3'
5'
C.cand5
C.cand4
C.cand3
C.cand2
C.cand1
lin-37

FIG. 20

[illegible]

lin-37 encodes a novel 32 kD hydrophilic protein

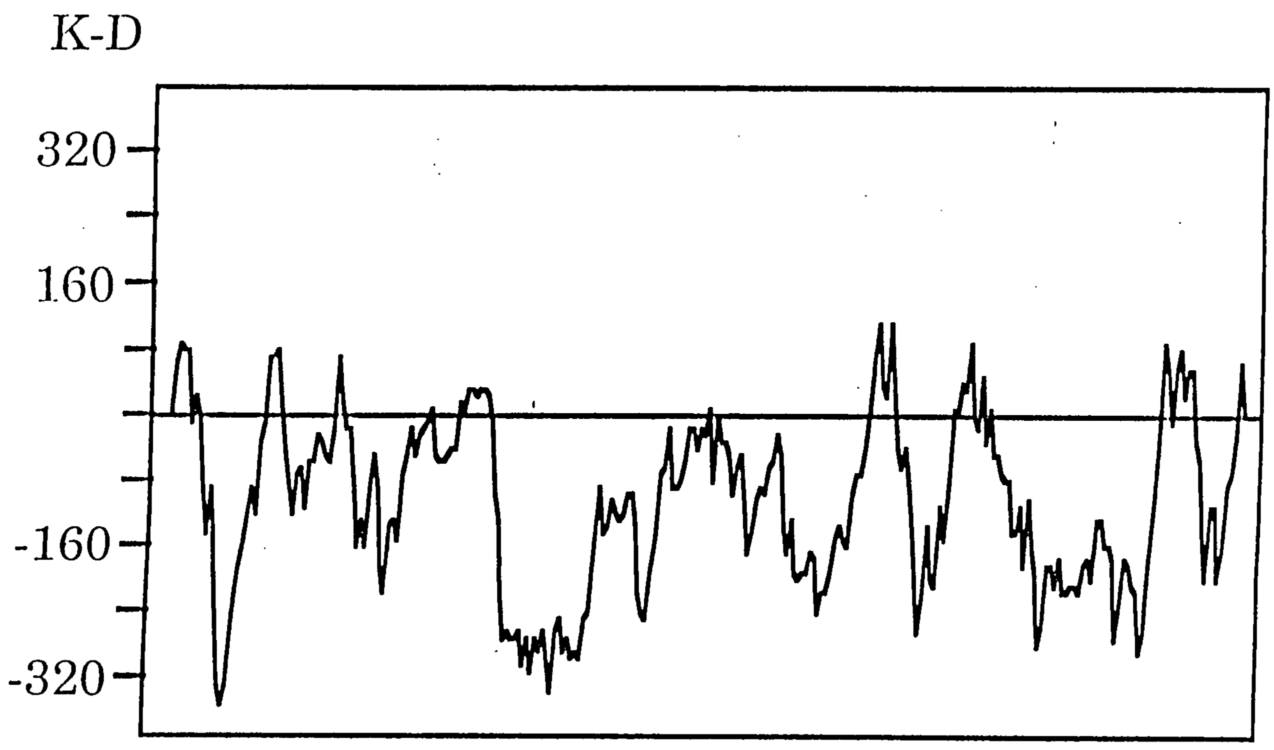
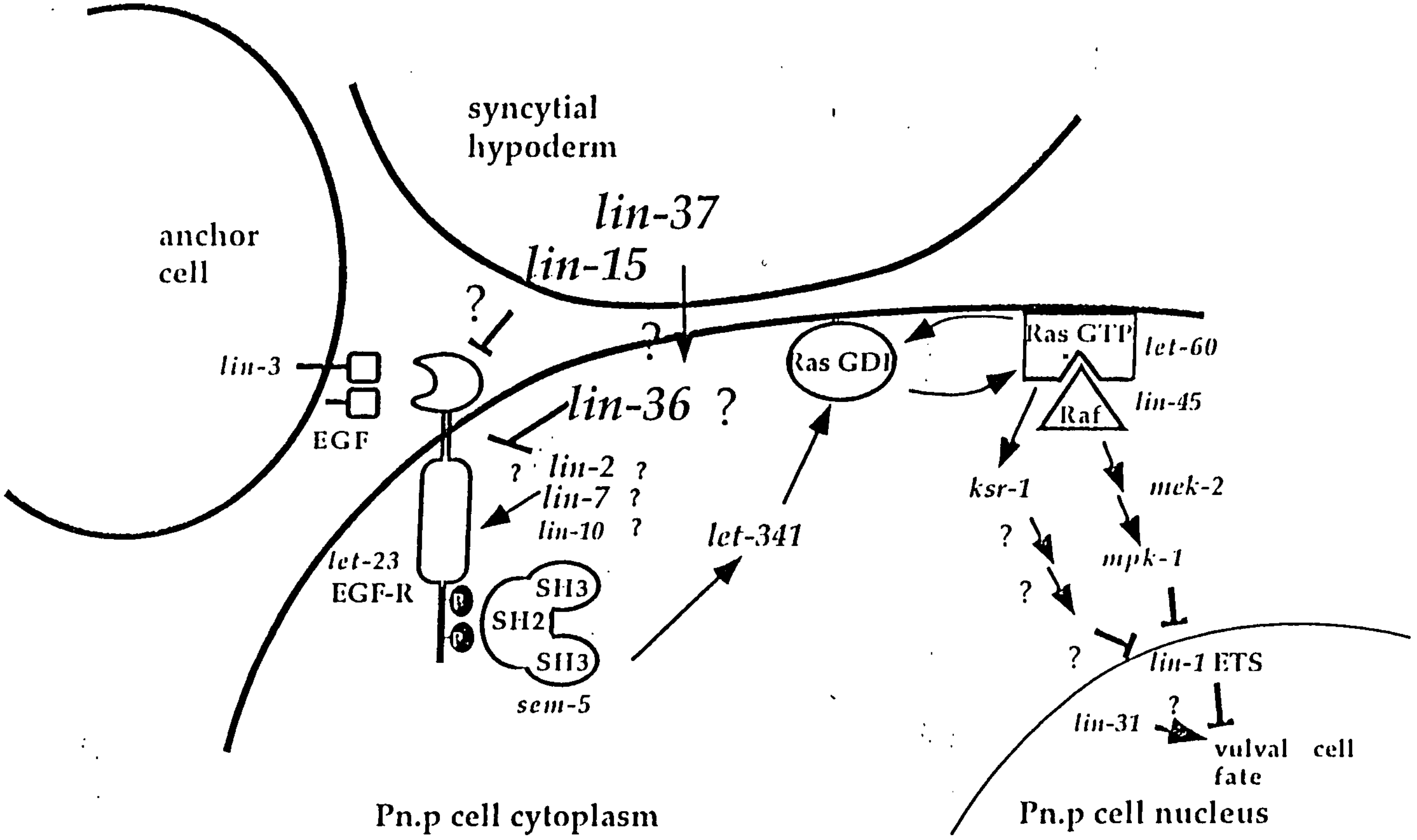


FIG. 21

A Model for *lin-37* function

lin-37 acts cell-nonautonomously, presumably in hyp7 (Hedgecock and Herman, Genetics **141**:989-1006). *lin-37* could act through *lin-36* to regulate localization of LET-23 in the Pn.p cells. Further expression and gene interaction studies on *lin-37* and other synMuv genes may shed some light on their function.



anchor cell
syncytial hypoderm
lin-37
lin-15
?
?
lin-3
EGF
lin-36 ?
? lin-2 ?
lin-7 ?
lin-10 ?
let-23
EGF-R
P
P
SH3
SH2
SH3
sem-5
let-341
Ras GDI
Ras GTP
let-60
Raf
lin-45
ksr-1
mek-2
?
?
mpk-1
?
lin-1 ETS
?
lin-31
vulval cell fate
Pn.p cell cytoplasm
Pn.p cell nucleus

FIG. 22

Results indicate that LIN-36 is able to self-associate. No other interactions among the synMuvs were detected.

		PREYS			
		lin-9	lin-15A	lin-15B	lin-36
	SNF1 (- control)	-	-	-	-
	lin-9	-	-	-	-
BAITS	lin-15A	-	-	-	-
	lin-15B	-	-	-	-
	lin-36	-	-	-	++

FIG. 23

Characterization of Molecular Interactions of the synMuv Gene Products

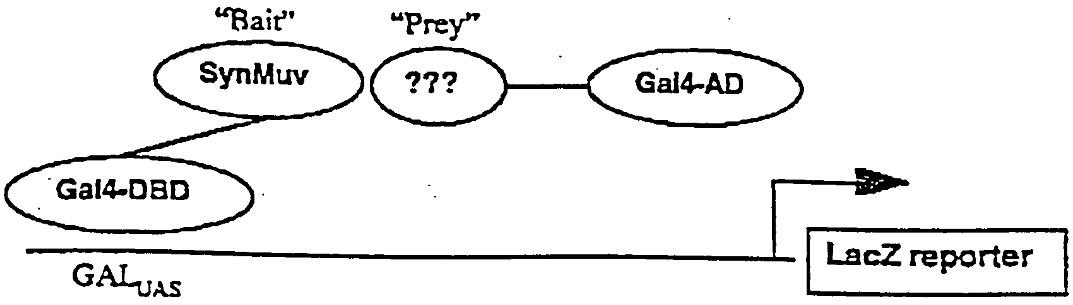
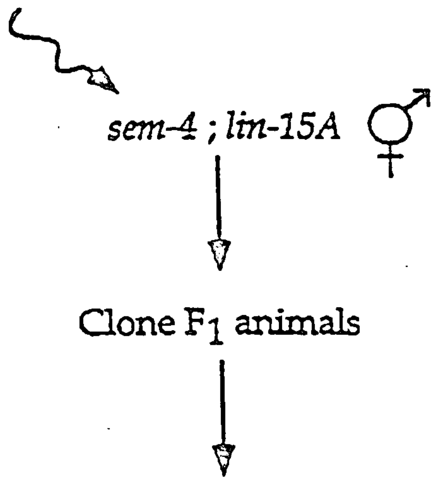

"Bait"
"Prey"
SynMuv
???
Gal4-AD
Gal4-DBD
LacZ reporter
GAL$_{UAS}$

FIG. 24

Future Search for Additional synMuv Alleles



sem-4 ; *lin-15A* ⚥

Clone F_1 animals

Inspect F_2 plates for Muv animals.
If sterile, and thus non-Egl, recover siblings from the F_2 plate.

FIG. 25

Sequence of *Mus musculus* cDNA clone vp19d01

1

GTGCTAAAACAATTGGTGAACCAGACTTGCCCAAAGTTGTTAATAT
TAACTTCTGTCCATCTGGTTTAAGGGTCTGATTGCCAAGTTTAAGAT
CAGATGTCTGTGATACTTTGTTTAAAATAATCTGATTGGCTGATATA
GTCACAGGAGTCTGAGCACCAAGTTTTTGAAGGCCACTTGGAAATG
CTGGTTTCACTGTGGTATTAGAATCTGCTTTAGAAACTGTGGTATTC
ACTGCAACTTGGTTAGTGTGGTTACTGTACACTGTGATTGGTTCCGT
GGAAATGGGCGTGGCTGTAGAGTCACCGGTAGAATTTATGTTGACA
ATTTCTTCCAGCTCTGTCTCCATGGGAATTGGGGATGACACAATTAC
AGCCTCAATACTATCCTCATCCACTAACGTTATAGCAGTGTCCATTA
TGTCGTCTGGAAGCAAACTATTCACTCGG

449 SEQ ID NO:15

FIG. 26

Homo sapiens lin-54-like cDNA Sequence

```
1
AAACTGGTCCAAGCTGAATTCCGGGGAAAAGGTGAAAATGGCGGA
TCTTTCGAAATACAATCCCGGCCCCTGACATACCAGAGGCGGCGGC
GACGGCGACGTCGCCAACCCGGCTCCCTTCTCGGTCCCCGGGTACC
CTGGAGCGCTCCAGTTTGGACAAACTGGGAAAGGAAGCTGCTGGC
CAGAGCTATCGTGCACGGGCAAAACAAGCCCCGCCGGGGCTCGGG
ATTGCCCAGGACCTTCCGGAGCCCCTACCTCGGAGCCCGAGGGAG
GAGAAAGCGGCAGCCGCTGGAGTGCCTGATCAACTTTCTAGCGGG
AGAACGATCATGGAGGTGGTGCCAGCTGAGGTGAATAGTTTGCTTC
CAGAGGAAATAATGGACACTGGTATAACTTTAGTGGATGATGATAG
TATTGAGGCTGTTATTGTTTCATCCCCAATTCCCATGGAGACAGAAC
TGGAAGAAATTGTCAACATAAATTCTACTGGTGACTCTACAGCCAC
GCCCATTTCCACGGAACCAATCACAGTGTACAGTAACCACACTAAC
CAAGTTGCAGTGAATACCACAATTACTAAAGCAGATTCTAATACCA
CAGTGAAACCAGCTTTTCCAAGTGGCCTTCAAAAACTTGGTGCTCA
GACTCCTGTGACTATATCAGCCAATCAGATTATTTTAAACAAAGTA
TCACAGACATCTGATCTTAAACTTGGCAATCAGACCCTTAAACCAG
ATGGACAGAAGTTAATTTTAACAACTTTGGGCAAGTCTGGTTCACC
AATTGTTTTAGCACTACCCCATAGCCAACTACCCCAGGCTCAGAAA
GTTACAACTCAGGCCCAGTCAGGAGATGCTAAGTTACCACCGCAGC
AAATTAAAGTAGTTACCATTGGAGGGAGGCCAGAGGTGAAACCTG
TCATTGGTGTCTCAGCATTGACCCCAGGAAGTCAACTGATTAATAC
TACAACTCAGCCCTCTGTGTTACAGACCCAACAGTTAAAAACAGTA
CAGATTGCTAAGAAGCCTCGAACGCCAACCTCTGGTCCAGTAATCA
CGAAGCTGATCTTTGCAAAACCAATTAATAGTAAAGCAGTTACAGG
ACAGACAACTCAAGTTTCACCACCAGTCATTGCAGGTAGGGTTCTT
TCACAGTCTACTCCCGGAACTCCATCAAAGACCATAACAATATCTG
AAAGTGGTGTTATTGGATCAACTTTAAATTCTACAACACAGACACC
AAATAAAATAGCCATCTCACCTTTGAAATCGCCAAATAAGGCAGT
GAAATCAACTGTGCAGACCATCACTGTTGGAGGAGTGAGCACATC
ACAGTTTAAGACAATTATTCCTCTGGCAACTGCTCCCAATGTCCAG
CAGATTCAAGTGCCTGGAAGCAAGTTTCATTATGTCCGACTTGTTAC
TGCCACATCAGCCAGTAGCTCAACCCAGCCAGTTAGTCAGAATCCC
AGTACAAACACTCAGCCTTTTCAGCAAGCAAAGCCAGTGGTTGTTA
ATACAACCCCAGTGCGGATGTCAGTTCCAATTGTCTCAGCTCAGGC
TGTCAAACAAGTTGTTCCAAAACCAATCAATCCAACTTCACAAATA
GTAACTACTAGCCAGCCACAGCAACGGCTTATCATGCCTGCCACAC
CACTGCCACAGATCCAGCCCAACCTCACTAACCTGCCACCAGGCA
CTGTCCTGGCACCAGCTCCGGGAACAGGGAATGTGGGTTATGCAGT
GCTTCCAGCTCAGTATGTTACTCAGCTACAGCAGTCTTCATATGTAT
CAATAGCAAGCAACTCTACCTTTACTGGAACATCTGGTATCCAGAC
CCAGGCACGGCTTCCATTCAATGGCATAATCCCATCAGAGTCGGCC
AGTCGGCCCCGAAAGCCCTGTAATTGTACAAAATCACTGTGTTTGA
AATTGTATTGTGATTGCTTTGCAAATGGTGAATTTTGCAACAACTGC
AATTGTACTAATTGTTACAACAATTTG
                              2000
```

FIG. 27 PAGE 1 OF 2

Homo sapiens lin-54-like cDNA Sequence, cont.

2001
GAACATGAAAATGAAAGGCAAAAAGCAATAAAGGCATGCCTTGA
CAGAAATCCAGAAGCCTTTAAGCCTAAGATAGGGAAAGGAAAGG
AGGGAGAATCTGATCGACGTCATAGCAAAGGGTGTAATTGCAAAC
GATCAGGATGTCTTAAAAACTACTGTGAATGCTATGAGGCAAAAAT
AATGTGTTCCTCAATATGCAAATGTATTGGCTGTAAGAATTTTGAAG
AAAGCCCGGAAAGGAAGACATTGATGCATTTGGCAGATGCAGCTG
AAGTAAGGGTACAGCAACAAACAGCAGCCAAGACGAAGTTATCCT
CTCAAATTTCAGACTTGCTTACTAGGCCAACACCAGCTTTAAATAGT
GGAGGCGGAAAATTGCCATTTACATTTGTAACTAAGGAANTANCTG
AAGCCACNTGTAATTGCCNCCTTGCCCAGGCAGAGCAGGCAGACA
AGAAGGGAAAATCAAAGGCAGCAGCGGAACGGATGATACTTGAG
GAATNCGGACNATGTTTGATGANTGTCATCAACTCTGCAGGAAAG
GCAAAAANTGACCCTTGTGCCCATGAATTGCTAACTCTTGCACAAA
AGACTGATAAATGGAACTGTACAGAAAATTTAAGGTGCAGGGACA
CTTGATTTTCTGGAAGAAAAACAATTACTGTATTTTAATTCAGTCCT
TGTTTTAAAAGACCTGAAATTATAATACTGAAGGAGAAGAAATTTT
AAATGAGGAAATTAGTACATTTTAAATCTTAGTTAAATCTGCTTATG
CCCCTTCTAAATTGAATTTTTCTNTATTATATAGATTTTTTAATTTGC
TTGGGTTTCTTAAGAATTAGATGTTCTCTTTCTGATACCTTTGACAAA
AAATGTTTATAAATTCATATAATTTATAATGTATGGTGTTGTATGAC
TTTGTTAATAGAAAAGCCAAAGCAGCAGTGGTTAGCACCCATTCTT
TGGGACTTGATCTAGAATATCTGCAGACAGAATGTTACATAAACAA
ATTCTTATGAAACACATTCAAATGACATTTTGTATTTAGAAAAGGA
CTATCTTTTAAAGAAAAAGCAGCCTTTTAGGGCCGATTCTGGAATA
ATATCCTGTTGTCACTTTGGGAATGTCAGAAGGGGAAACAATCCCC
AGGCACACTAAAAATTTTTTAAAGTTATTTAAAAAAACATATAAAA
TATTAAAGGACAGTAAATCTCAGAGGATGGGCAATGTGTTTCTATA
ATAAGGAAAGGCTAACAGATGCTCTGGGCTGTCTCCATTTTCTTTCA
AAGAGGTGGTATGTATTTGAAGTAATAAATTGTCAAAGTGATTACT
GGGTACTATTAAAATGATAGGTGGATATAAATGGAAGTAAACATTA
TGTAGTGATAATATAGAACCTCACATAGTAATCAAGTATAAAATTT
GGCATGGGTGGAGAAACAAAGNATCGGGAAGCTGCCAAAGATGA
ATTTAGAGAAGTTTTCATCTATACAATCAATTATTTTTACAGACTTTT
TTTCCGGAATTCTTTTGCT

3535 **SEQ ID NO:16**

FIG. 27 PAGE 2 OF 2